JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Foliage Diagnostics: Pioneering Insights through Feature Extraction Techniques in Leaf Disease Identification

[1]Asha Patil
Shrimad Rajchandra Institute of Management and Computer Application, Uka Tarsadia University, Bardoli, Gujarat, India

Abstract: For thousands of years, agriculture has been the main source of human sustenance. Approximately 50% of the global population still depends on it for their subsistence. Every year, plant leaf diseases cause significant losses in crop productivity across the globe. To offset the financial losses resulting from plant leaf diseases, it is imperative to maintain the health of the plants at different phases of their growth and development. Since plant leaves are the primary organs that display symptoms of illnesses, leaves are frequently utilized to detect and identify diseases. Visual observation is a difficult method of disease detection that takes a great deal of human skill. Digital image processing techniques and computational intelligence offer enhanced assistance to farmers in identifying leaf diseases. Plant leaf diseases can be identified by their symptoms, which can be extracted as features. Therefore, in these kinds of systems, feature extraction techniques are essential. The review of feature extraction approaches are highlighted in the paper. It offers a thorough analysis of numerous visual characteristics, including colour, texture, and shape, for a range of illnesses in diverse cultural contexts.

IndexTerms **- Digital Image processing, Leaf diseases, Image Feature extraction.**

I. Introduction

Agriculture perform a crucial role in sustaining livelihoods in India, with a majorshare of the population directly or indirectly engaged in this sector. Farmers strive for improved crop quality and productivity by carefully managing factors such as temperature, light, and humidity[1]. Given the challenges posed by population growth and climate variations, the agricultural industry seeks innovative approaches to enhance food production. However, identifying diseases in plant leaves remains a persistent challenge for farmers. Naked-eye observation is not always reliable, prompting the exploration of automated expert systems for timely disease detection[2]. Utilizing machine learning algorithms in image processing methods provides efficient solutions, despite the challenges associated with extracting and selecting features from plant leaf images. Fundamental features such as color, texture, and shape play a crucial role in achieving accurate identification [3]. The primary objective of this initiative is to develop a system capable of identifying plant diseases, integrating advanced digital image processing and ML techniques for enhanced accuracy and efficiency and comparing and evaluating various feature extraction methodess.

The current paper is structured into five sections. Section II delves into the literature review. III defines numerous feature extraction techniques. Section IV equates different feature extraction techniques. Section V conclude the work with future research directions.

II. Previous work

To ensure the robust growth of the apple industry, it is imperative to employ precise and swift disease detectors for altering leaf spots, , grey spots, brown spots, mosaics, and rust apple disease [4].The turmeric plant leaf are mainly affected by diseases like Leaf Spot and Leaf Blotch for that texture analysis involves examining pixel relationships in a local area of leaf images. A statistical method for texture analysis that account the spatial connection of pixels is the Grey Level Co-occurrence Matrix. By identifying pixel pairs with specific values in a particular orientation and distance, GLCM describes the textural properties of an image and calculates its statistical features[5]. Plant leaf disease analysis done by Gabor filters. Features including the radial centre frequency, standard deviation, and orientation make up the Gabor filters. In order to avoid dimensionality problems, Gabor filters must be reduced huge feature sizes[6]. In recognizing leaf disease images, the Local Binary Pattern (LBP) employs the value of the center pixel as a threshold for the 3x3 neighbouring pixels. The threshold process yields a binary pattern that represents the textural qualities.[7].The crutial features arefundamentallyreducingsize of feature set which denote the remarkable portions of unhealthy region of leaf. SIFT offers a bunch of features of an leaf image[8]. Wheat leaf diseases include Alternaria, Anthracnose, bacterial spot, canker, and others. The color model is made up of three elements: L represents the luminosity layer, and a*b represents the chromaticity layer. It is essential to properly identify ailments in wheat leaf [9].Feature extraction technique reduce feature data set while keeping the important parts. When classifying these features data set, it depend on pattern recognition. The image features usually contain colour, shape and texture features. Currently, most of the researchers target plant leaf texture as the most important feature for disease classification of leaf plant[10].The primary characteristic defining the overall leaf structure is the curvature along the leaf edge. In the initial application of contour-based feature extraction for identifying leaf diseases, various shape properties of leaf images, including eccentricity, extent, perimeter, solidity, equivalent diameter, and perimeter area ratio, were employed [11][12].

III. FEATURE EXTRACTION TECHNIQUES

This section explores and evaluates various strategies for feature extraction. Different facets of disease identification are investigated using image processing and machine learning, either independently or in conjunction.[13]. The feature extraction techniques are as below:

3.1Texture-based Features

When there is significant tonal variation in a tiny area of an image, the texture dominates that area. It essentially depicts the spatial layouts and colour patterns found in the picture. Texture perception is influenced by light, contrast, distance, and direction. Entropy, contrast, Skewness, variance, homogeneity, and other factors can all be used to describe the texture of an image. This section reviews the primary methods of feature extraction for texture description.

3.1.1Gray level co-occurrence matrices

Gray-level co-occurrence matrices serve as a statistical method for deriving second-order measures that depict texture. The characteristics unveiled through this process are known as Haralick features. The GLCM takes the form of a 2-D square matrix with an order of N, where N corresponds to the number of grey levels within an image. This matrix illuminates the relationship between pixels, revealing how many pixels with a (value of i) exist at a specific distance from pixels with a (value of j). Essentially, the characteristics of each pixel are intricately determined by the direction and distance in this interplay. The GLCM matrix is calculated at four angles—0°, 45°, 90°, and 135°—capturing. The fundamental concept underlying pixel values lies in their role in classifying leaf diseases from images through the computed GLCM characteristics. By examining individual pixels, the image features of Septoria leaf blight, downy mildew, and frog eye in soybean crops provide a valuable indicator of the diversity in patterns of pixel values within the image. This examination involves understanding the relationship between a pixel and its neighboring pixels at specific distances and directions.

3.1.2 Local Binary Pattern

A grayscale visual descriptor called local binary pattern (LBP) is used to quantify binary patterns inside a circular area. The correlation between each and every pixels inside a neighborhood is provided. The local spatial details of a picture are defined by this non-parametric operator. The LBP operator thresholds the center pixel value and labels every pixel in the neighborhood with one of two binary values: 0 or 1 [14]. The label is 1 if there is a positive difference in the grey levels between the center pixel and a neighboring pixel, and 0 otherwise. The threshold binary pattern is multiplied by the weight of the corresponding pixels, and the total of these multiplications is used to determine the LBP code. The LBP code, represented as a histogram, can be regarded as a feature in and of itself[15].

3.1.3 Scale-invariant feature transform

Local essential features of objects remain unaltered even when they undergo scale transformations according to the scale-invariant feature transform (SIFT) technique. From any given image, a set of local key points is extracted to create a feature vector. Moreover, object matching makes use of the SIFT approach. To validate this similarity, the Euclidean distance between the keypoints in two images is calculated. Subsequently, the features of the newly generated image are compared with the keypoints obtained from the reference image[16]. Three primary processes are involved in extracting local features from an image: recognising scale-space maxima and keypoints, orientation, and defining the keypoint descriptor. With SIFT, a tiny area can offer a vast collection of important points. The SIFT characteristics establish a local region that remains invariant to variations in noise, light, direction and size. These traits are reasonably simple to extract, however there is a problem with large dimensionality, which increases computing complexity [17].

3.1.4 Gabor filters

Inspired by the human visual system, the Gabor filter method is applicable on local texture aspects of images. On the behalf of Gabor wavelet and Gaussian window, Gabor filters offer features at various sizes and angle. The Gaussian function plays a role in determining the window size in the Gabor wavelet, which, in turn, constitutes a windowed rapid Fourier transform. The generalized Gabor wavelet form can be used to define the 2-D Gabor filter.[18]. The study tried to identify illnesses like Alternaria, bacterial blight, and anthracnose by using Gabor filters to extract the uni-chrome properties from photos of pomegranate plant leaves. Gabor filters were used with LBP and Haralick features in another Apple illness diagnosis algorithm. In order to increase the identification rate, prior research on soybean diseases used Gabor and colour features. Additionally, a research study that was proposed indicated that the Gabor transform is a reliable and efficient method for texture analysis [19]. Additionally, use the AdaBoost classifier to detect canker infection in citrus plants, Gabor features are extracted at six scales and eight rotations. [20].

3.1.5 Speed-up robust features:

A new rotation and scale invariant feature descriptor called speed-up robust features (SURF) was inspired in part by the SIFT descriptor. In terms of computation, reproducibility, robustness, and uniqueness, this method is superior to SIFT. This technique is working in finding of plenat leaf disease identification because of its advantage in feature extraction. In order to identify different illnesses, researchers created a set of SURF properties that were taken from photos of cassava crop leaves.[21]. Infections including grey leaf spot , northern leaf blight, and common rust in maize crops are also studied using the SURF approach. [22].

3.2 Colour-based Features

Colour features reflect sensor response for various wavelengths and give colours their physical and visual characteristics. Colour features are relatively stable to direction and scale and resilient against complicated backdrops. Photometrical data such as lighting, shadowing, shading, and optical density of colour channels are provided by colour characteristics. As was covered in previous sections, the colour features can be expressed in a variety of colour spaces, including HSI, Luv, RGB, HSV, L*a*b* and YCbCr. One can just utilise the grayscale values in each band as features[23].The other feature measures could be obtained from the colour spaces covered in below:

3.2.1 Color histogram

An image's colour distribution and brightness and contrast can be described using the colour histogram. A colour histogram can be thought of as a collection of bins that show the likelihood that a given pixel is a certain colour.[24]. It focuses on colour composition regardless of colour placement or spatial organisation. You can see a colour histogram as the vector shown below:

$$H = \{Ho, H1, H2, H3 \dots \dots \dots .. Hn\} \qquad (1)$$

In order to identify different plant species, the leaf disease image calculated colour histograms for the R, G, and B channels. After examining many bins, the optimal result was obtained using 10 histogram bins [25].

3.2.2 Color moments:

Colour moments, which are independent of scale and rotation, show how comparable colours are in a picture. As a result, it can be used in applications for image recognition and retrieval. Probability theory served as the inspiration for the idea of colour moments, which are used to describe colour attributes and uniquely express the probability distribution[26].Therefore, if a probability distribution is used to describe the colors, moments can be used to understand the colour distribution. Three or more channels are used to encode colors. Moments are determined for every one of these channels. As a result, nine moments three for each colour band are used to describe the image with three colour bands. The mean, Skewness, and standard deviation are the three color moments utilized to economically and effectively depict the color distribution of photographs. Assuming there are N total pixels in the image and that a pixel f_{ij} indicates the i^{th} color band at the j^{th} pixel.[27].

3.3 Shape-based Features:

Shape features, which include perimeter borders, circular, triangular, and rectangular shapes, give an object in an image its visual characteristics. Properties including rotation, identifiability scale invariance, translation, and statistical independence, are followed by shape feature extraction algorithms. The definitions of several shape feature descriptors are as follows[28].

1. Center of Gravity, also known as the centroid, is calculated for both region-based and contour-based form descriptions. It provides the average coordinate of the pixels in an image.
2. Eccentricity:The measure of eccentricity is defined as the ratio of the major axis length to the minor axis length. Another name for it is a circularity ratio, which is a line aspect ratio measurement that ranges from 0 to 1 for circular spots, respectively.
3. Orientation: The estimated angle _ between the principal axis of the spot and the horizontal is known as orientation.
4. Aspect Ratio (AR): This articulates the relationship between the width and height of the image or the selected region and is represented using the mathematical symbolic notation x: y.
5. Area (a): The area is calculated for each region or image within the range of N to M.
6. Rectangularity (Extent): It is a measurement of the spot's shape's rectangularity. Its value is also in the range of 0 and 1. If the rectangularity value of a location is 1, then its shape is perfectly rectangular.
7. Ratio of principle axis: It can be defined as the ratio of the spot's main to minor axis lengths.
8. Euler Number: It illustrates the relationship between a spot's surrounding areas and hole count.
9. Perimeter: The spot perimeter's shape was computed.

3.4 Combination of feature extraction techniques:

KSE-100 index is an index of 100 companies selected from 580 companies on the basis of sector leading and market the analysis of plant diseases also makes use of a variety of features. It is possible to mix various feature types to enhance the efficiency of disease detection applications. These pairings may be helpful if various feature kinds offer complementing data [29][30].

3.5 Result and Discussion:

The dataset and attributes taken into consideration might affect the choice of image analysis method and how well it performs.
Table-1 Comparison of feature extraction method accuracy

Extraction Method	Features	Dataset/ Plant Leaf	Accuracy
Region-based[11]	Shape	Flavia datasets	75.5 %
GLCM[5]	Texture	-	91%
GLCM[26]	Color	-	92%
LBP[31]	Texture	Corn Leaf	81.1%.
Gabor filters[19]	Texture +color	Tomattow	90.37%.
GLCM[29]	Texture +color	PlantVillage	98.79%
LBP[15]	Color+ Shape +Texture	Repositories Digi-pathos	98.5%

Higher accuracy is frequently achieved by combining combination of colour, shape, and texture. Selecting the right features is essential depending on the features of the dataset and the particular goals of the identitying the leaf disease from image.

3.6 Conclusion:

The topic of this article describes many processes such as capturing images, processing, identifying key features, selecting features, and determining the type of disease. The plant's shape, texture, and colour are examples of its features. Systems that use one or more of these characteristics are developed to determine whether a plant leaf is diseased or healthy. It was shown that combining several aspects produces superior outcomes than utilizing only one. Plant infections and diseases typically progress in stages. It would be intriguing to assess how well feature extraction methods work at various phases of the leaf disease. Early detection could assist farmers in reducing loss. It is also possible to do more studies to assess various feature extraction methods and how they work together to identify various plant diseases. Differentiating between several diseases is easier than identifying only one. A few scientists have experimented with combining different characteristics to find plant infections. More work is needed to investigate how different features might be combined for improved accuracy.

REFERENCES:

[1] B. V. Nikith, N. K. S. Keerthan, M. S. Praneeth, and D. T. Amrita, "Leaf Disease Detection and Classification," *Procedia Comput. Sci.*, vol. 218, pp. 291–300, 2022, doi: 10.1016/j.procs.2023.01.011.

[2] A. A. Ahmed and G. Harshavardhan Reddy, "A Mobile-Based System for Detecting Plant Leaf Diseases Using Deep Learning," *AgriEngineering*, vol. 3, no. 3, pp. 478–493, 2021, doi: 10.3390/agriengineering3030032.

[3] L. Li, S. Zhang, and B. Wang, "Plant Disease Detection and Classification by Deep Learning - A Review," *IEEE Access*, vol. 9, no. Ccv, pp. 56683–56698, 2021, doi: 10.1109/ACCESS.2021.3069646.

[4] K. P. Singh, N. Basant, and S. Gupta, "Support vector machines in water quality management," *Anal. Chim. Acta*, vol. 703, no. 2, pp. 152–162, 2011, doi: 10.1016/j.aca.2011.07.027.

[5] G. Kuricheti and P. Supriya, "Computer vision based turmeric leaf disease detection and classification: A step to smart agriculture," *Proc. Int. Conf. Trends Electron. Informatics, ICOEI 2019*, no. Icoei, pp. 545–549, 2019, doi: 10.1109/ICOEI.2019.8862706.

[6] M. K. R. Gavhale and P. U. Gawande, "An Overview of the Research on Plant Leaves Disease detection using Image Processing Techniques," *IOSR J. Comput. Eng.*, vol. 16, no. 1, pp. 10–16, 2014, doi: 10.9790/0661-16151016.

[7] P. Ouppaphan, "Corn Disease Identification from Leaf Images Using Convolutional Neural Networks," *ICSEC 2017 - 21st Int. Comput. Sci. Eng. Conf. 2017, Proceeding*, vol. 6, pp. 233–238, 2018, doi: 10.1109/ICSEC.2017.8443919.

[8] D. Angayarkanni and L. Jayasimman, "Leaf Disease Recognition Using Segmentation With Visual Feature Descriptor," pp. 2624–2629, 2022, doi: 10.21917/ijivp.2022.0372.

[9] S. Nema and A. Dixit, "G. Kuricheti and P. Supriya, 'Computer vision based turmeric leaf disease detection and classification: A step to smart agriculture,' Proc. Int. Conf. Trends Electron. Informatics, ICOEI 2019, no. Icoei, pp. 545–549, 2019, doi: 10.1109/ICOEI.2019.8862706.," *2018 Int. Conf. Circuits Syst. Digit. Enterp. Technol. ICCSDET 2018*, pp. 1–5, 2018, doi: 10.1109/ICCSDET.2018.8821098.

[10] K. Anitha and S. Srinivasan, "Feature Extraction and Classification of Plant Leaf Diseases Using Deep Learning Techniques," *Comput. Mater. Contin.*, vol. 73, no. 1, pp. 233–247, 2022, doi: 10.32604/cmc.2022.026542.

[11] D. P. Mahurkar and H. Patidar, "Revealing leaf species through specific contour and region-based features extraction," *e-Prime - Adv. Electr. Eng. Electron. Energy*, vol. 5, no. August, p. 100228, 2023, doi: 10.1016/j.prime.2023.100228.

[12] A. Patil and K. Lad, "Chili Plant Leaf Disease Detection Using SVM and KNN Classification," *Adv. Intell. Syst. Comput.*vol. 1187, pp. 223–231, 2021, doi: 10.1007/978-981-15-6014-9_26.

[13] V. K. Vishnoi, K. Kumar, and B. Kumar, *A comprehensive study of feature extraction techniques for plant leaf disease detection*, vol. 81, no. 1. Springer US, 2022.

[14] S. Devi and B. Gupta, "GLCM-LBP Plant Leaf Disease Detection," *Int. J. Sci. Res. Eng. Dev.*, vol. 2, no. 3, pp. 135–141, 2019, [Online]. Available: www.ijsred.com.

[15] M. Ashfaq, W. Ahmad, S. M. Adnan Shah, and M. Ilyas, "Citrus Leaf Disease Detection using Joint Scale Local Binary Pattern," no. January, 2020.

[16] G. Wang and J. Wang, "SIFT Based Vein Recognition Models: Analysis and Improvement," *Comput. Math. Methods Med.*, vol. 2017, 2017, doi: 10.1155/2017/2373818.

[17] H. S. Jayswal and J. P. Chaudhari, "Plant Leaf Disease Detection and Classification using Conventional Machine Learning and Deep Learning," vol. 11, no. 3, pp. 1094–1102, 2020.

[18] E. Miao, G. Zhou, and S. Zhao, “Research on Soybean Disease Identification Method Based on Deep Learning,” *Mob. Inf. Syst.*, vol. 2022, 2022, doi: 10.1155/2022/1952936.

[19] Wiharto, F. H. Nashrullah, E. Suryani, U. Salamah, N. P. T. Prakisya, and S. Setyawan, “Texture-based feature extraction using gabor filters to detect diseases of tomato leaves,” *Rev. d’Intelligence Artif.*, vol. 35, no. 4, pp. 331–339, 2021, doi: 10.18280/ria.350408.

[20] L. Shen and L. Bai, “AdaBoost Gabor feature selection for classification,” *Image Vis. Comput. Conf.*, pp. 77–83, 2004, [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.95.4078&rep=rep1&type=pdf.

[21] J. S. H. Albayati and B. B. Üstündağ, “Evolutionary feature optimization for plant leaf disease detection by deep neural networks,” *Int. J. Comput. Intell. Syst.*, vol. 13, no. 1, 2020, doi: 10.2991/ijcis.d.200108.001.

[22] W. Albattah, M. Nawaz, A. Javed, M. Masood, and S. Albahli, “A novel deep learning method for detection and classification of plant diseases,” *Complex Intell. Syst.*, vol. 8, no. 1, pp.507–524, 2022, doi: 10.1007/s40747-021-00536-1.

[23] J. M. M. Araujo and Z. M. A. Peixoto, “A new proposal for automatic identification of multiple soybean diseases,” *Comput. Electron. Agric.*, vol. 167, no. October, p. 105060, 2019, doi: 10.1016/j.compag.2019.105060.

[24] H. Ali, M. I. Lali, M. Z. Nawaz, M. Sharif, and B. A. Saleem, “Symptom based automated detection of citrus diseases using color histogram and textural descriptors,” *Comput. Electron. Agric.*, vol. 138, pp. 92–104, 2017, doi: 10.1016/j.compag.2017.04.008.

[25] A. Caglayan, O. Guclu, and A. B. Can, “A plant recognition approach using shape and color features in leaf images,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8157 LNCS, no. PART 2, pp. 161–170, 2013, doi: 10.1007/978-3-642-41184-7_17.

[26] M. K. DEVI, “Plant Leaf Disease Detection,” *Interantional J. Sci. Res. Eng. Manag.*, vol. 07, no. 08, pp. 55–62, 2023, doi: 10.55041/ijsrem25396.

[27] J. H. Dewan and S. D. Thepade, “Image Retrieval Using Low Level and Local Features Contents: A Comprehensive Review,” *Appl. Comput. Intell. Soft Comput.*, vol. 2020, 2020, doi: 10.1155/2020/8851931.

[28] J. Kainat, S. Sajid Ullah, F. S. Alharithi, R. Alroobaea, S. Hussain, and S. Nazir, “Blended Features Classification of Leaf-Based Cucumber Disease Using Image Processing Techniques,” *Complexity*, vol. 2021, 2021, doi: 10.1155/2021/9736179.

[29] N. Ahmed, H. M. S. Asif, and G. Saleem, “Leaf Image-based Plant Disease Identification using Color and Texture Features,” no. February, 2021, doi: 10.1007/s11277-021-09054-2.

[30] N. Ahmad, H. M. S. Asif, G. Saleem, M. U. Younus, S. Anwar, and M. R. Anjum, “Leaf Image-Based Plant Disease Identification Using Color and Texture Features,” *Wirel. Pers. Commun.*, vol. 121, no. 2, pp. 1139–1168, 2021, doi: 10.1007/s11277-021-09054-2.

[31] A. Rachmad, M. Syarief, S. Rifka, F. Sonata, W. Setiawan, and E. M. S. Rochman, “Corn Leaf Disease Classification Using Local Binary Patterns (LBP) Feature Extraction,” *J. Phys. Conf. Ser.*, vol. 2406, no. 1, 2022, doi: 10.1088/1742-6596/2406/1/012020.